OM protein - protein search, using sw model

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

Run on: February 17, 2005, 15:19:48 ; Search time 0.001 Seconds
(without alignments)
258.063 Million cell updates/sec

Title:
Perfect score: 2623
Sequence: 1 MDPKLGRMAASLLAVLLLLLL.........NYIEGTKLFAAFPLEMAQLH 507

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 509 residues

Total number of hits satisfying chosen parameters: 1

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   aay76144.geneseqp2000s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result   %   Query   Match Length DB   ID   Description
No.   Score   Match   ----------   -----------

1   2612.5   99.6   509   1   AAY76144   Human secreted pro

ALIGNMENTS

RESULT 1
AAY76144
ID   AAY76144 standard; protein; 509 AA.
XX
AC   AAY76144;
XX
DT   23-MAR-2000   (first entry)
XX
DE   Human secreted protein encoded by gene 21.
XX
KW   Human; secreted protein; cancer; tumour; developmental abnormality;
KW   foetal deficiency; blood disorder; immune system disorder; inflammation;
KW   autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW   schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW   atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW   digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
KW   therapy; chromosome 18q22-23.
XX
OS   Homo sapiens.
XX
PN   WO995860-A1.
XX
PD   18-NOV-1999.
XX
PF   06-MAY-1999;   99WO-US009847.
XX
PR   12-MAY-1998;   98US-0085093P.
PR   12-MAY-1998;   98US-0085094P.
PR   12-MAY-1998;   98US-0085105P.
PR   12-MAY-1998;   98US-0085180P.
PR   18-MAY-1998;   98US-0085906P.
PR   18-MAY-1998;   98US-0085920P.
PR   18-MAY-1998;   98US-0085921P.
PR   18-MAY-1998;   98US-0085922P.
PR   18-MAY-1998;   98US-0085924P.
PR   18-MAY-1998;   98US-0085925P.
PR   18-MAY-1998;   98US-0085927P.
PR   18-MAY-1998;   98US-0085928P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
PI   Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR, Lafleur DW;
PI   Endress GA, Ebner R;
XX
DR   WPI; 2000-062296/05.
DR   N-PSDB; AAZ65270.
XX
PT   New isolated human genes and the secreted polypeptides they encode,
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders.
XX
PS   Claim 11; Page 373-374; 475pp; English.
XX
CC   AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
CC   AAY76124 to AAY76223 are the secreted proteins encoded by the 97 human
CC   genes. The gene encoding this protein was found to be on chromosome 18q22
CC   -23. The genes and their corresponding secreted polypeptides are useful
CC   for preventing, treating or ameliorating medical conditions, e.g. by
CC   protein or gene therapy. Also pathological conditions can be diagnosed by
CC   determining the amount of the new polypeptides in a sample or by
CC   determining the presence of mutations in the new genes. Specific uses are
CC   described for each of the 97 genes, based on which tissues they are most
CC   highly expressed in, and include developing products for the diagnosis or
CC   treatment of cancer, tumours, developmental abnormalities and foetal
CC   deficiencies, blood disorders, diseases of the immune system, autoimmune
CC   diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
CC   schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
CC   atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
CC   digestive/endocrine disorders, infections and AIDS. The polypeptides are
CC   also useful for identifying their binding partners. The sequences shown
CC   in AAY76224 to AAY76424 represent fragments of the secreted proteins
XX
SQ   Sequence 509 AA;

Query Match   99.6%;   Score 2612.5;   DB 1;   Length 509;
Best Local Similarity   99.8%;   Pred. No. 0;   Matches 507;   Conservative 0;   Mismatches 0;   Indels 1;   Gaps 1;

Qy   1 MDPKLGRMAASLLAV-LLLLLLERGMFSSPSPPPPALLEKVFQYIDLHQDEFVQTLKEWVAI 59
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MDPKLGRMAASLLAVLLLLLLLERGMFSSPSPPPPALLEKVFQYIDLHQDEFVQTLKEWVAI 60

Qy   60 ESDSVQPVPRFRQELFRMWAVAADTLQRLGARVASVDMGPQOLPDGQSLPIPPVILAELG 119
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ESDSVQPVPRFRQELFRMWAVAADTLQRLGARVASVDMGPQOLPDGQSLPIPPVILAELG 120

Qy   120 SDPTKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSA 179
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 SDPTKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSA 180

Qy   180 FRALEQDLPVNIKFIIEGMEEAGSVALEELVEKXDRFSGVDYIVISDNLWISQRKPAI 239
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 FRALEQDLPVNIKFIIEGMERAGSVALEELVEKXDRFSGVDYIVISDNLWISQRKPAI 240

Qy   240 TYGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEV 299
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TYGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEV 300

```
Qy   300  VPLTEEEINTYKAIHLDLEEYRNSSRVEKFLPDTKEEIIMHLWRYPSLSIHGIEGAFDEP  359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  VPLTEEEINTYKAIHLDLEEYRNSSRVEKFLPDTKEEIIMHLWRYPSLSIHGIEGAFDEP  360

Qy   360  GTKTVIPGRVIGKFSIRLVPHNMVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWI  419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GTKTVIPGRVIGKFSIRLVPHNMVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWI  420

Qy   420  ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHS  479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHS  480

Qy   480  QNEKINRWNYIEGTKLFAAFFLEMAQLH  507
          ||||||||||||||||||||||||||||
Db   481  QNEKINRWNYIEGTKLFAAFFLEMAQLH  508
```

Search completed: February 17, 2005, 15:19:48
Job time : 0.001 secs

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: February 17, 2005, 15:19:48 ; Search time 0.001 Seconds
                 (without alignments)
                 258.063 Million cell updates/sec

Title: us-10-036-342-57
Sequence: 1 MDPKLGRMAASLLAVLLLLL.........NYIEGTKLFAAFFLEMAQLH 507
Perfect score: 2623

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 509 residues

Total number of hits satisfying chosen parameters: 1

                 Minimum DB seq length: 0
                 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   aay76144.geneseqp2000s:*

                 Pred. No. is the number of results predicted by chance to have a
                 score greater than or equal to the score of the result being printed,
                 and is derived by analysis of the total score distribution.

                              SUMMARIES

Result   Query   Match Length DB   ID
No.   Score   %   DB   ID   Description
                 Match   ----------
1   2612.5 99.6   509   1   AAY76144   Human secreted pro

                              ALIGNMENTS

RESULT 1
AAY76144
ID   AAY76144 standard; protein; 509 AA.
XX
AC   AAY76144;
XX
DT   23-MAR-2000 (first entry)
XX
DE   Human secreted protein encoded by gene 21.
XX
KW   Human; secreted protein; cancer; tumour; developmental abnormality;
KW   foetal deficiency; blood disorder; immune system disorder; inflammation;
KW   autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW   schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW   atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW   digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
KW   therapy; chromosome 18q22-23.
OS   Homo sapiens.
XX
PN   WO9958660-A1.
XX
PD   18-NOV-1999.
XX
PF   06-MAY-1999;   99WO-US009847.
XX
PR   12-MAY-1998;   98US-0085093P.

PR   12-MAY-1998;   98US-0085094P.
PR   12-MAY-1998;   98US-0085105P.
PR   12-MAY-1998;   98US-0085180P.
PR   18-MAY-1998;   98US-0085906P.
PR   18-MAY-1998;   98US-0085920P.
PR   18-MAY-1998;   98US-0085921P.
PR   18-MAY-1998;   98US-0085922P.
PR   18-MAY-1998;   98US-0085923P.
PR   18-MAY-1998;   98US-0085924P.
PR   18-MAY-1998;   98US-0085925P.
PR   18-MAY-1998;   98US-0085927P.
PR   18-MAY-1998;   98US-0085928P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
PI   Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR, Lafleur DW;
PI   Endress GA, Ebner R;
XX
DR   WPI; 2000-0622296/05.
DR   N-PSDB; AAZ65270.
XX
PT   New isolated human genes and the secreted polypeptides they encode,
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders.
XX
PS   Claim 11; Page 373-374; 475pp; English.
XX
CC   AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
CC   AAY76124 to AAY76223 are the secreted proteins encoded by the 97 human
CC   genes. The gene encoding this protein was found to be on chromosome 18q22
CC   -23. The genes and their corresponding secreted polypeptides are useful
CC   for preventing, treating or ameliorating medical conditions, e.g. by
CC   protein or gene therapy. Also pathological conditions can be diagnosed by
CC   determining the amount of the new polypeptides in a sample or by
CC   determining the presence of mutations in the new genes. Specific uses are
CC   described for each of the 97 genes, based on which tissues they are most
CC   highly expressed in, and include developing products for the diagnosis or
CC   treatment of cancer, tumours, developmental abnormalities and foetal
CC   deficiencies, blood disorders, diseases of the immune system, autoimmune
CC   diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
CC   schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
CC   atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
CC   digestive/endocrine disorders, infections and AIDS. The polypeptides are
CC   also useful for identifying their binding partners. The sequences shown
CC   in AAY76224 to AAY76424 represent fragments of the secreted proteins
SQ   Sequence 509 AA;
XX

Query Match   99.6%; Score 2612.5; DB 1; Length 509;
Best Local Similarity   99.8%;
Matches 507; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy   1 MDPKLGRMAASLLAV-LLLLLLERGM 25
              ||||||||||||||| ||||||||||
Db   1 MDPKLGRMAASLLAVLLLLLLERGM 25

Qy   26 ESDSVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELG 119
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ESDSVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELG 120

Qy   120 SDPTKGTVCFYGHLLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSA 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 SDPTKGTVCFYGHLLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSA 180

Qy   180 FRALEQDLPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAI 239
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 FRALEQDLPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAI 240

Qy   240 TYGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEV 299
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TYGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEV 300

```
Qy    300  VPLTEEEINTYKAIHLDLEEYRNSSRVEKFLPDTKEEILMGHLWRYPSLSIHGIEGAPDEP  359
           ||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db    301  VPLTEEEINTYKAIHLDLEEYRNSSRVBKFLPDTKEEILMHLWRYPSLSIHGIEGAPDEP   360

Qy    360  GTKTVIPGRVIGKFSIRLJVPHMNVSAVEKQVTRHLBDVFSKRNSSNKMVVSMTLGLHPWI  419
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  GTKTVIPGRVIGKFSIRLJVPHMNVSAVEKQVTRHLBDVFSKRNSSNKMVVSMTLGLHPWI  420

Qy    420  ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSSTIPIAKMFQBIVHKSVVLIPLGAVDDGEHS  479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSSTIPIAKMFQBIVHKSVVLIPLGAVDDGEHS  480

Qy    480  QNEKINRWNYIEGTKLPAAFFLEMAQLH  507
           ||||||||||||||||||||||||||||
Db    481  QNEKINRWNYIEGTKLPAAFFLEMAQLH  508
```

Search completed: February 17, 2005, 15:19:48
Job time : 0.001 secs

```
OM protein - protein search, using sw model

                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:             February 17, 2005, 15:23:26 ; Search time 0.001 Seconds
                                               (without alignments)
                                               161.226 Million cell updates/sec

Scoring table:      us-10-036-342-57
Sequence:           1 MDPKLGRMAASLIAVLLLLL..........NYIEGTKIFAAPFLBMAQLH 507
Perfect score:      2623
Title:

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5

Searched:           1 seqs, 318 residues

Total number of hits satisfying chosen parameters:     1

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing:    Minimum Match 0%
                    Maximum Match 100%
                    Listing first 45 summaries

Database :          aay08759.geneseqp1990s:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.


                            SUMMARIES

Result        %       Query
  No.   Score Match Length DB  ID              Description
  1     176   6.7   318  1  AAY08759           L. monocytogenes d


                            ALIGNMENTS

RESULT 1
AAY08759
ID    AAY08759 standard; protein; 318 AA.
XX
AC    AAY08759;
XX
DT    13-AUG-1999  (first entry)
XX
DE    L. monocytogenes dapE protein.
XX
OS    Listeria monocytogenes.
XX
PN    WO9929884-A2.
XX
PD    17-JUN-1999.
XX
PF    11-DEC-1998;   98WO-EP008096.
XX
PR    11-DEC-1997;   97DB-01054938.
XX
PA    (VEIC/) VON EICHEL-STREIBER C.
PA    (CHAK/) CHAKRABORTY T.
XX
PI    Von Eichel- Streiber C, Chakraborty T;
XX
DR    WPI; 1999-371353/31.
DR    N-PSDB; AAX77872.
XX
PT    A targeted genetic conditioning used in, e.g. gene therapy.
XX
PS    Claim 11; Page 55-57; 58pp; German.
XX
CC    This invention describes a novel bacterium modified to release a foreign
CC    DNA under the control of eukaryotic regulatory elements into a foreign
CC    organism, organ, tissue, cell line or individual cells after infection of
CC    the host. The method of the invention allows a targeted genetic
CC    conditioning (TGC) for inducing targeted somatic transgenesis in an
CC    animal host where a bacteria comprising a foreign DNA integrated into an
CC    episomal vector releases, in the case of infection of a foreign organism,
CC    organ, tissue, cell line or individual cells, the foreign DNA under the
CC    control of eukaryotic regulatory elements for later transcription and/or
CC    expression, causing transcription and expression of foreign DNA and/or
CC    foreign protein in the location. The method is used for induced somatic
CC    gene therapy, immunological protection against microbes or protection
CC    against tumours. Listeria monocytogenes can be modified so that it has
CC    reduced pathogenicity in animals and can produce foreign proteins that
CC    are then secreted into, e.g. milk. Examples of where the method are
CC    useful, include treatment of cystic fibrosis (by manipulation of an
CC    airborne bacteria), treatment of beta-thalassemia by ex vivo modification
CC    of a haematopoietic cell to contain beta-globulin, for treatment of
CC    Hurler syndrome by infection of CD34+ cells of the bone marrow with an
CC    alpha-L-iduronidase gene and also to treat Fanconi's anaemia with a group
CC    C complementation gene. The method does not require genetic
CC    transformation/modification of the organism requiring gene therapy. The
CC    method allows for safe and ethical gene therapy in an organism. The
CC    expression of a therapeutic foreign protein by a bacteria that infects an
CC    animal is advantageous in that the expression can be tissue- or organ-
CC    specific
XX
SQ    Sequence 318 AA;

Query Match             6.7%;  Score 176;  DB 1;  Length 318;
Best Local Similarity   22.4%;
  Matches   83;  Conservative  55;  Mismatches 158;  Indels  74;  Gaps  12;

Qy    127 VCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSAFRALEQD 186
              :||:||      |      :    |   |  || |      :  ||      ::
Db      4 LAFSGHMDVVDAGDVSKWKFPPEATEHSGKLYGRGATDMKSGLAAMVIAMIBLHEKQK 63

Qy    187 LPVNIKFIIEGMEEAGSVALEBLVEKEKDRFFSGVDYIVISDNIWISQRKPA---ITYGT 243
              |      |  ||      |  :||           |  | |  |    |      ||
Db     64 LNGKIRLLATVGEEICBLGAEQLTQK-----GYADDLHGLIIGE-----PSGHRIVYAH 112

Qy    244 RGNSYFMVEVCRCDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEVVPLT 303
              || |   |   :|    || || |   |      | ||    :    |   |     :
Db    113 KGSINY--FVKSTGKNAHSS----------------------------------MP 132

Qy    304 EEBINTYKAIHLDLEEYRNSSRVEKFL--PDTKEEILMHLWRYPSLSIHGIEGAPDEPGT 361
              ||      :     |    | |    |      || || |   :  || | | |   |
Db    133 ESGVNAIDNLLLFYNE-----VEKFVKSVDATNEILGPF------IHNVT-VIDGGNQ 178

Qy    362 KTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLBDVFSKRNSSNKMVSMTLGLHPWIAN 421
              |   |     |  | :     |   ||  :  :||   |    |   |  |  :|
Db    179 VNSIPEKAQLQGNIRSIPEMDNETV-KQVLVKIINKLNKQSNVNLELI-PDYDKQPVFSD 236

Qy    422 IDDTQYLAAKRAIRTVFGTEEPDMIRDGSTIPIAKMFQBIVHKSVVLIPLGAVDDGEHSON 481
              |  |  |   | :|   |       | :||   | |     : |  || ||  | |  :
Db    237 KNSDLVHIAKSVASDIVKEEIPLLGISGTTDAASFTK--AKKEFPVIIFGPGNETPHQVN 294

Qy    482 EKINRWNYIE 491
              :||   ||:|
Db    295 ENVSIGNYLE 304
```

OM protein - protein search, using sw model

Run on: February 17, 2005, 16:02:03 ; Search time 0.001 Seconds
                                              (without alignments)
                                              125.229 Million cell updates/sec

Title:
Perfect score: 2623
Sequence: 1 MDPKLGRMAASILIAVLLLL............NYIEGTKLFAAFFLEMAQLH 507

Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5
Searched: 1 seqs, 247 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   aaw19339.geneseqp1990s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result   Query
No.   Score Match Length DB ID   Description
---- ----- ----- ------ -- -------------- --------------------
1   57   2.2   247   1   AAW19339   Hybrid polypeptide

                             ALIGNMENTS

RESULT 1
AAW19339
ID   AAW19339 standard; protein; 247 AA.
XX
AC   AAW19339;
XX
DT   17-OCT-2003 (revised)
DT   25-MAR-2003 (revised)
DT   11-MAR-1998 (first entry)
XX
DE   Hybrid polypeptide P102 for detecting antibodies to HIV-1.
XX
KW   Hybrid; chimeric; fusion; polypeptide; P102; detection; antibody;
KW   VCPM B-5876; pol; beta-galactosidase; reverse transcriptase;
KW   immunodetection; infection.
XX
OS   Human immunodeficiency virus 1.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   RU2071502-C1.
XX
PD   10-JAN-1997.
XX
PF   28-FEB-1992; 92SU-05029950.
XX
PR   28-FEB-1992; 92SU-05029950.

XX
PA   (BIOS=) BIOSERVIS BIOTECHN CO.
XX
PI   Zaitsev IZ, Sukhanova LL, Alatortseva GI;
XX
PR   WPI; 1997-371403/34.
DR   N-PSDB; AAT73713.
XX
PT   Hybrid polypeptide P102 for detecting antibodies to HIV-1 - comprises HIV
PT   -1 reverse transcriptase fused to E.coli beta-galactosidase.
XX
PS   Claim 1; Col 11; 6pp; Russian.
XX
CC   The present sequence is the hybrid polypeptide P102 for detecting
CC   antibodies to human immunodeficiency virus type 1 (HIV-1). P102 is
CC   produced by E. coli VCPM B-5876 transformed by recombinant DNA plasmid
CC   pHP102 carrying a fragment of the HIV-1 pol gene fused to the E. coli
CC   beta-galactosidase gene. P102 is capable of binding antibodies to HIV-1
CC   reverse transcriptase (i.e. the pol gene product). P102 is useful for
CC   immunodetection of HIV-1 infection. (Updated on 25-MAR-2003 to correct PF
CC   field.) (Updated on 25-MAR-2003 to correct PR field.) (Updated on 17-OCT-
CC   2003 to standardise OS field)
XX
SQ   Sequence 247 AA;

Query Match   2.2%; Score 57; DB 1; Length 247;
Best Local Similarity   23.5%; Pred. No. 0;
Matches   32; Conservative 18; Mismatches 56; Indels 30; Gaps 5;

Qy   275 DLVALLGSLLVDSSGHILVPGI------------YDEVVPLTEEEINTYKAIHDLEEY 320
              |||||||:||:| ||:||||||||||
Db   114 DIQKLVGKLNMASQ--IYPGIKVRQLCKLLRGTKALIEVVPLTEEA-------LELAEN 163

Qy   321 RN--SSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAPDEPGTKTVIPGRVIGKFSIRLV 378
              | :|:|||:||||:|||||| |: | :||||||||::|| |:| :|||||||:|
Db   164 REILKEPVHGVYYDPSKDLIAEIQKQG----QGPEAALQDSGLEVNIITDSQYALGIQAQ 219

Qy   379 PHMNVSAVEKQVTRHL 394
              || || :||:|:|
Db   220 PDKSESELVNQIIEQL 235

Search completed: February 17, 2005, 16:02:03
Job time : 0.001 secs

us-10-036-342-57

OM protein - protein search, using sw model

Run on: February 17, 2005, 16:05:18 ; Search time 0.001 Seconds

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Title: us-10-036-342-57
Perfect score: 2623
Sequence: 1 MDPKLGRMAASLLAVLLLLL..........NYIEGTKLFAAFFLEMAQLH 507
                                  (without alignments)
                                  135.369 Million cell updates/sec

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5
Searched: 1 seqs, 267 residues
Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   aar40171.geneseqp1990s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	33	1.3	267	1	AAR40171	Sequence encoded b

ALIGNMENTS

RESULT 1
AAR40171

ID   AAR40171 standard; protein; 267 AA.
XX
AC   AAR40171;
XX
DT   08-FEB-1994 (first entry)
XX
DE   Sequence encoded by chicken brain cDNA encoding aceylcholine receptor-
DE   inducing activity (ARIA) protein E18.
XX
KW   acetylcholine receptor-inducing activity protein; ARIA; E18.
XX
OS   Gallus.
XX
PN   US5237056-A.
XX
PD   17-AUG-1993.
XX
PF   29-MAY-1991; 91US-00706872.
XX
PR   29-MAY-1991; 91US-00706872.
XX
PA   (HARD) HARVARD COLLEGE.
XX
PI   Fischbach GD;
XX
DR   WPI; 1993-272184/34.
DR   N-PSDB; AAQ46887.
XX
PT   DNA encoding protein which co-purifies with acetyl:choline receptor-
PT   inducing activity - used to obtain prods. for treating deficiencies in
PT   neural transmitter receptors, e.g. Alzheimer's disease.
XX
PS   Disclosure; Fig 2; 17pp; English.
XX
CC   ARIA promotes the synthesis and accumulation of acetylcholine receptors
CC   (AChR8) in muscle cells. The protein was isolated on the basis of its
CC   ability to increase the rate of insertion of AChRs into the surface
CC   membrane of chick myotubes. Oligos with sequences corresp. to the
CC   sequence of the protein were used to amplify from chicken brain cDNA a 34
CC   -nucleotide sequence encoding 11 AAs from ARIA. The 34-nucleotide
CC   sequence corresp. to the 3' nucleotides of the 7th AA codon through the
CC   3' nucleotide of the 18th AA codon. Each primer contd. an EcoRI site at
CC   its 5' end. The first pair of oligos are AAQ46882/Q46883. The second pair
CC   are AAQ46884/Q46885. The oligos are nested pairs. The PCR amplified prod.
CC   was used as a probe to screen an E18 chick brain cDNA library. AAQ46887
CC   is the insert from a positive clone obtd. by the screening an E18 chick
CC   brain cDNA library in lambda gt10 and introduced into a plasmid
CC   designated p65-21
XX
SQ   Sequence 267 AA;

Query Match   1.3%;   Score 33; DB 1; Length 267;
Best Local Similarity   75.0%;   Pred. No. 0;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy   9 AASLLAVLLLLL 20
           |:||||||||||
Db   249 ADTWLAVLLLLL 260

Search completed: February 17, 2005, 16:05:18
Job time : 0.001 secs

Run on: February 17, 2005, 16:08:02

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Title: us-10-036-342-57
Perfect score: 2623
Sequence: 1 MDPKLGRMAASLLAVLLLLL.........NYIEGTKLFLFAAFFLEMAQLH 507

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 906 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : aar64250.geneseqp1990s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
                    %
Result        Query Match
No.    Score   Match Length DB  ID              Description
------ ------- ----- ------ --  --------------  ---------------
  1     55     2.1    906    1  AAR64250        Human mGluR1B. Ho
```

ALIGNMENTS

RESULT 1
AAR64250

```
ID   AAR64250 standard; protein;  906 AA.
XX
AC   AAR64250;
XX
DT   22-DEC-1994  (revised)
DT   21-JUL-1995  (first entry)
XX
DE   Human mGluR1B.
XX
KW   Metabotropic glutamate receptor; mGluR1; mGluR1B; cerebellum; pCMV-T7-3.
OS   Homo sapiens.
XX
PN   WO9429449-A1.
XX
PD   22-DEC-1994.
XX
PF   03-JUN-1994;  94WO-US006273.
PR   04-JUN-1993;  93US-00072574.
XX
PA   (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.
PI   Daggett L, Ellis SB, Liaw C, Ponteler A, Johnson EC, Hess SD;
XX
DR   WPI; 1995-036478/05.
DR   N-PSDB; AAQ80416.
XX
PT   New DNA encoding human metabotropic glutamate receptor sub:type(s) - and
PT   related proteins, probes, RNA, transformed cells and antibodies, useful
PT   in treatment, diagnosis and identification of specific modulators.
XX
PS   Claim 3; Page 75-79; 125pp; English.
XX
CC   A human cerebellum cDNA library was screened with DNA encoding the rat
CC   mGluR1 receptor. Several clones were isolated, and a complete sequence
CC   was obtained by ligating overlapping clones. The resulting construct
CC   (given in AaQ80416) encoded human mGluR1B (AAR64250). The encoding
CC   sequence was operatively linked to the regulatory elements in pCMV-T7-3
CC   for expression in mammalian cells. (Updated on 25-MAR-2003 to correct PN
CC   field.)
XX
SQ   Sequence 906 AA;
```

```
Query Match             2.1%;  Score 55;  DB 1;  Length 906;
Best Local Similarity   17.2%;  Pred. No. 0;
Matches   80;  Conservative  58;  Mismatches  143;  Indels  184;  Gaps  22;

Qy     102 LPDGQSLP------------------------------------IPPVILAELGSDPTKGTVVCFY 130
               ||||||||
Db     141 LPDGQSLPPGRTKKPIAGVIGPGSSSVAIQVQNLLQLFDIPQIAYSATSIDLSDKTLYKY 200

Qy     131 GHLDVQPAD------------RGDGWLTDPYLTEVDGKLYGRGATD------NKGPVLA 172
               :: :    ::                   :::                    :
Db     201 -FLRVVPSDTLQARAMLDIVKRYNWTYVSAVHTEGN--YGESGMDAFKELAAQEGLCIA 256

Qy     173 WINAVSA------------FRALEQDLP-----------VNIKFIEGMEEAGSVALEELVE 211
                 :::            :::  :                   :: :     :
Db     257 HSDKIYSNAGEKSFDRLLRKLRERLPKARVVVCFCEGMTVRGLLSAMRRLGVVGEFSLIG 316

Qy     212 K-----EKDRFFSGVDYIVISDNLWISQRKPAITYGTRGNSYFM---VEVKCRD---QDFH 261
               :          : :                            : :
Db     317 SDGWADRDEVIEGYE-VEANGGITIKLQSPEV--RSFDDYFLKRLDTNTRNPWFPEFW 372

Qy     262 SGTFP-----GGILHEP-----------------MADLVALLGSLVDSSGHILV 292
               ::         ::::                         :: ::  :
Db     373 QHRFQCRLPGHLLENPNFKRICTGNESLEENYVQDSKMGFVINAIYAMAHGLQNMHALC 432

Qy     293 P---GIYDEVVPLTEEEINTYKAIHDLEEYRNSSRVEKFLPDTKEEILMHLWRYPSLSI 349
                    :: :: : :                           :::          ::
Db     433 PGHVGLCDAMKPI-----------------DGSKLLDFLIKSS------FIGV 462

Qy     350 HGIEGAFDEPGTKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVV 409
               : :  ::       :        :  :                :    :
Db     463 SGEEVWFDEKGD---APGR----YDI------MNLQYTE-------ANRYDY 494

Qy     410 SMTLGLHPWIANIDDTQYLAAKRAIRTVFGTEP-----DMIRDG 448
               : :  : :   :::: ::                    :::::
Db     495 VHVGTWHEGVLNIDDYKIQMNKSGVVRSVCSEPCLKGQIKVIRKG 539
```

Search completed: February 17, 2005, 16:08:02
Job time : 0.001 secs

PT   Search time 0.001 Seconds
     (without alignments)
     459.342 Million cell updates/sec

OM protein - protein search, using sw model

Run on: February 17, 2005, 16:11:01 ; Search time 0.001 Seconds

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

Title: U8-10-036-342-57
Perfect score: 2623
Sequence: 1 MDPKLGRMAASILAVLLLL.........NYIEGTKLFAAFFLEMAQLH 507
                                                     (without alignments)
                                                     497.874 Million cell updates/sec

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 982 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%;
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   aar09301.geneseqp1990s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Match Length	DB ID	Description
1	63.5	2.4	982	1 AAR09301	Sequence deduced f

ALIGNMENTS

RESULT 1
AAR09301

Query Match   2.4%; Score 63.5; DB 1; Length 982;
Best Local Similarity 20.3%;
Matches   44; Conservative 33; Mismatches 67; Indels 73; Gaps 13;

Qy   298 EVVPLTEEEINTYKAIHLDLEEYRN--SSRVEKFLFDTKEEILMHL------WRYPSLS 348
          |||||||||| :::::   :
Db   425 EVVPLTEE-------AELELAENREILKEPVHGVYYDPSKDLIAELQKQGDGQWTYQ-- 474

Qy   349 IHGIEGAFDEP--GTKTVIPGRVIGKFSIRLVPHMN-----VSAVEKQVTRHLEDVFSKR 401
          :   : :   :   :: :   :   :
Db   475 ------IYQBPFKNLKT-------GKYARTRGAHTNDVKQLTEAVQKIATESI--VIWGK 519

Qy   402 NSSNKMVVSM----TLGLHEWIAN-IDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKM 456
          :   :   : :   : :   :: :   : :
Db   520 TPKFKLPIQKETWETWWIEYWQATWIPEWEFV-----------------NTPPLVKL 559

Qy   457 FQEIVHKSVVLIPLGAVD---DGEHSQNEKINRWNYI 490
          : :   :   :: : :: :
Db   560 WYQLEKEPII----GAETFYVDGAANRETKLGKAGYV 592

OM protein - protein search, using sw model

```
                         GenCore version 5.1.6
                         Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:                  February 17, 2005, 16:17:31 ; Search time 0.001 Seconds
                                              (without alignments)
                                              51.207 Million cell updates/sec

Title:           us-10-036-342-57
Perfect score:   2623
Sequence:        1 MDPKLGRMAASLLAVLLLL..........NYIEGTKLFAAFFLEMAQLH 507

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1 seqs, 101 residues

Total number of hits satisfying chosen parameters:    1

Database :       aar91823.genpept:*

Post-processing: Minimum Match 0%
Maximum DB seq length: 2000000000
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

                      *
Result           Query
No.     Score   Match Length DB   ID            Description
------------------------------------------------------------
1       38.5    1.5   101    1    AAR91823
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1
AAR91823

Query Match   1.5%;   Score 38.5;   DB 1;   Length 101;
Best Local Similarity 25.3%; Pred. No. 0;
Matches   19; Conservative 10; Mismatches 25; Indels 21; Gaps   4;

ALIGNMENTS

```
Qy   196 EGMEEAGSVALEELVEKEKDRFF-----SGVDYIVISDNLWISQRKPA-----------ITY 241
             :|||.|         :||:|.|         :||                       :|:
Db     4 QSLEESGG----DLVKPEGSLTITCTASGFSF---STNWICWVRQAPGKGLEWIACIDA 56

Qy   242 GTRGNSYFMVEVKCR 256
         |:|:.:|:|:.:|:
Db    57 GSSGTTYYASWAKGR 71
```

Search completed: February 17, 2005, 16:17:31
Job time : 0.001 secs

```
OM protein - protein search, using sw model
                                        GenCore version 5.1.6
                                        Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:             February 17, 2005, 16:19:15 ; Search time 0.001 Seconds
                                        (without alignments)
                                        7.605 Million cell updates/sec

Title:              us-10-036-342-57
Perfect score:      2623
Sequence:           1 MDPKLGRMAASLIAVLLLLL...........NYIEGTKLFAAFFLEMAQLH 507
Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched:           1 seqs, 15 residues
Total number of hits satisfying chosen parameters:   1

Database :          aaw39029.geneseqp1990s.*
                    Listing first 45 summaries

Post-processing:    Minimum Match 0%
                    Maximum Match 100%
Minimum DB seq length: 0
Maximum DB seq length: 200000000

                                        SUMMARIES

Result   Score   Query            %
No.              Match  Length  DB  ID             Description
-------  -----   -----  ------  --  --------       ------------------
1        40      1.5    15      1   AAW39029       Peptide resembling

                                        ALIGNMENTS

Result 1
AAW39029

ID    AAW39029 standard; peptide; 15 AA.
XX
AC    AAW39029;
XX
DT    27-MAR-1998  (first entry)
XX
DE    Peptide resembling an SH3 domain binding peptide SEQ ID NO:428.
XX
KW    Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW    tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW    PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS    Synthetic.
XX
PN    WO9730074-A1.
XX
PD    21-AUG-1997.
XX
PF    14-FEB-1997;  97WO-US002298.
XX
PR    16-FEB-1996;  96US-0060299.
XX
PA    (CYTO-) CYTOGEN CORP.
PA    (UYNC-) UNIV NORTH CAROLINA.
XX
PI    Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ,  Fowlkes DM;
PI    Rider JE;
XX
DR    WPI; 1997-424972/39.
XX
PT    Src homology region 3 binding peptide - used to activate Src tyrosine
PT    kinase(s) and to stimulate immune response by increasing production of
PT    certain lymphokine(s), e.g. interleukin-1.
XX
PS    Claim 22; Page 93; 131pp; English.
XX
CC    The present sequence represents a peptide which resembles a Src homology
CC    region 3 (SH3) binding peptide. SH3 binding peptides are selected from:
CC    (a) peptides which bind the SH3 domain of Cortactin; (b) peptides which
CC    bind the middle SH3 domain of Abl; (c) peptides which bind the SH3 domain
CC    of Abl; (d) peptides which bind the SH3 domain of Src; (e) peptides which
CC    bind the SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain
CC    of p53bp2; (g) peptides which bind the amino-terminal SH3 domain of Crk;
CC    (h) peptides which bind the SH3 domain of Yes; and (i) peptides which
CC    bind the amino-terminal SH3 domain of Grb2. The purified binding peptides
CC    can be used in the method to identify inhibitors of their binding to
CC    their respective SH3 domains, which could be used to modulate the
CC    pharmacological activity of proteins or polypeptide containing the SH3
CC    domain. The peptides can also be used to activate Src or Src-related
CC    protein tyrosine kinases, to stimulate the immune response by increasing
CC    the production of certain lymphokines, e.g. tumour necrosis factor-alpha
CC    and interleukin-1, or to deliver a conjugated molecule to certain
CC    cellular compartments containing Src or Src related proteins
XX
SQ    Sequence 15 AA;

Qy          26 SSPSPPP 32
               |||||||
Db           1 SSPSPPP 7

Query Match             1.5%;  Score 40;  DB 1;  Length 15;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

Search completed: February 17, 2005, 16:19:15
Job time : 0.001 secs

```
                                                              GenCore version 5.1.6
                                                              Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:              February 17, 2005, 17:29:59 ; Search time 0.001 Seconds
                                                  (without alignments)
                                                  257.556 Million cell updates/sec

Title:
Perfect score:       2623
Sequence:            1 MDPKLGRMAASLLAV................NYIEGTKLFAAFFLEMAQLH 507
                                                  us-10-036-342-57

Scoring table:       BLOSUM62
                     Gapop 10.0 , Gapext 0.5

Searched:            1 seqs, 508 residues

Total number of hits satisfying chosen parameters:        1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :           us-09-731-872-242:*

                     Pred. No. is the number of results predicted by chance to have a
                     score greater than or equal to the score of the result being printed,
                     and is derived by analysis of the total score distribution.


                                              SUMMARIES

              %
Result        Query
No.    Score  Match  Length DB  ID                             Description
------ -----  -----  ------ --- --------                       -----------
  1   2612.5  99.6    508   1   us-09-731-872-242


                                              ALIGNMENTS

RESULT 1
us-09-731-872-242

Query Match        99.6%;
Best Local Similarity   99.8%;
Matches   507;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   1;

                                                Score 2612.5;  DB 1;  Length 508;
                                                Pred. No. 0;

Qy      1 MDPKLGRMAASLLAV-LLLLLERGMFSSPSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAI 59
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db      1 MDPKLGRMAASLLAVLLLLLLERGMFSSPSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAI 60

Qy     60 ESDSVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQOLPDGQSLPIPPVILAELG 119
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ESDSVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQOLPDGQSLPIPPVILAELG 120

Qy    120 SDPTKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSA 179
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 SDPTKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSA 180

Qy    180 FRALEQDLPVNIKFIIEGMEEAGSVALLELVEKEKDRFFSGVDYIVISDNLWISQRKPAI 239
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 FRALEQDLPVNIKFIIEGMEEAGSVALLELVEKEKDRFFSGVDYIVISDNLWISQRKPAI 240

Qy    240 TYGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLIVALLGSLVDSSGHILVPGIYDEV 299
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TYGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLIVALLGSLVDSSGHILVPGIYDEV 300

Qy    300 VPLTEBEINTYKAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAFPDEP 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 VPLTEBEINTYKAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAFPDEP 360

Qy    360 GTKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWI 419
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GTKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWI 420

Qy    420 ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHS 479
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHS 480

Qy    480 QNEKINRWNYIEGTKLFAAFFLEMAQLH 507
          ||||||||||||||||||||||||||||
Db    481 QNEKINRWNYIEGTKLFAAFFLEMAQLH 508
```

Search completed: February 17, 2005, 17:30:00
Job time : 1 secs

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: February 17, 2005, 17:32:29 ; Search time 0.001 Seconds
           (without alignments)
           257.556 Million cell updates/sec

Title: us-10-036-342-57
Perfect score: 2623
Sequence: 1 MDPKLGRMAASLLAVLLLL......NYIEGTKLFAAFFLEMAQLH 507

Scoring table: BLOSUM62
           Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 508 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : us-09-948-783-139:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result   Query
No.   Score   Match Length DB ID   Description
---- ------ ----------- --   --   -----------
1   2612.5   99.6   508   1 us-09-948-783-139

ALIGNMENTS

RESULT 1
us-09-948-783-139

Query Match   99.6%; Score 2612.5; DB 1; Length 508;
Best Local Similarity   99.8%; Pred. No. 0;
Matches 507; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy   1 MDPKLGRMAASLLAV-LLLLLERGMFSSSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAI 59
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   1 MDPKLGRMAASLLAVLLLLLLERGMFSSSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAI 60

Qy   60 ESDSVQPVRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELG 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ESDSVQPVRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELG 120

Qy 120 SDPTKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSA 179
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 SDPTKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSA 180

Qy 180 FRALEQDLPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAI 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 FRALEQDLPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAI 240

Qy 240 TYGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEV 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 TYGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEV 300

Qy 300 VPLTEEEINTYKAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAFDEP 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 VPLTEEEINTYKAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAFDEP 360

Qy 360 GTKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNRMVVSMTLGLHPWI 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 GTKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNRMVVSMTLGLHPWI 420

Qy 420 ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLLIPLGAVDDGEHS 479
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db 421 ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLLIPLGAVDDGEHS 480

Qy 480 QNEKINRWNYIEGTKLFAAFFLEMAQLH 507
          ||||||||||||||||||||||||||||
Db 481 QNEKINRWNYIEGTKLFAAFFLEMAQLH 508

Search completed: February 17, 2005, 17:32:29
Job time : 0.001 secs

```
OM protein - protein search, using sw model

                GenCore version 5.1.6
                Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:         February 17, 2005, 17:33:32 ; Search time 0.001 Seconds
                                              (without alignments)
                                              258.063 Million cell updates/sec

Title:          us-10-036-342-57
Perfect score:  2623
Sequence:       1 MDPKLGRMAASLLAVLLLLL..........NYIEGTKLFAAFFLEMAQLH 507

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Database :      us-09-892-877-137:*

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Total number of hits satisfying chosen parameters:          1

Searched:       1 seqs, 509 residues

Scoring table:  us-09-892-877-137;

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.


                                       SUMMARIES

                     %          Query
Result               Match Length DB  ID
No.   Score          DB                         Description
----  -----          ------------     --------------------------------------
1     2612.5   99.6       509   1  us-09-892-877-137


                                      ALIGNMENTS

RESULT 1
us-09-892-877-137

Query Match        99.6%;        Score 2612.5;  DB 1;  Length 509;
Best Local Similarity  99.8%;    Pred. No. 0;
Matches   507;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   1;

Qy      1 MDPKLGRMAASLLAV-LLLLLERGMFSSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAI  59
          |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db      1 MDPKLGRMAASLLAVLLLLLLERGMFSSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAI  60

Qy     60 ESDSVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELG 119
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ESDSVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELG 120

Qy    120 SDPTKGTVCEYGHLDVOPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSA 179
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 SDPTKGTVCEYGHLDVOPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSA 180

Qy    180 FRALEQDLPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLMWISQRKPAI 239
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 FRALEQDLPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLMWISQRKPAI 240

Qy    240 TYGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHIIVPGIYDEV 299
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TYGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHIIVPGIYDEV 300

Qy    300 VPLTEEEINTYKAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAFDEP 359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 VPLTEEEINTYKAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAFDEP 360

Qy    360 GTKIVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWI 419
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GTKIVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWI 420

Qy    420 ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKWFQEIVHKSVVLIPLGAVDDGEHS 479
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKWFQEIVHKSVVLIPLGAVDDGEHS 480

Qy    480 QNEKINRWNYIEGTKLFAAFFLEMAQLH 507
          ||||||||||||||||||||||||||||
Db    481 QNEKINRWNYIEGTKLFAAFFLEMAQLH 508


Search completed: February 17, 2005, 17:33:32
Job time : 0.001 secs
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

Run on: February 17, 2005, 17:34:38 ; Search time 0.001 Seconds
                                            (without alignments)
                                            254.007 Million cell updates/sec

Title: us-10-036-342-57
Perfect score: 2623
Sequence: 1 MDPKLGRMAASLLAVLLLLL..............NYIEGTKLFAAFFLEMAQLH 507

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Database :   us-09-791-389-2:*

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Total number of hits satisfying chosen parameters: 1

Searched: 1 seqs, 501 residues

Scoring table: BLOSUM62

Sequence: 1 MDPKLGRMAASLLAVLLLLL..............NYIEGTKLFAAFFLEMAQLH 507

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

                    %
Result   Query
No.   Score   Match   Length DB   ID   Description
------ -----   -----   ------ --   --   -----------
1   2566.5 97.8   501   1   us-09-791-389-2

ALIGNMENTS

RESULT 1
us-09-791-389-2

Query Match   97.8%;
Best Local Similarity   99.4%;   Pred. No. 0;
Matches 498; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy   8 MAASLLAV-LLLLLERGMFSSPSPPPPALLEKVFQYIDLHQDEFVQTLKEWVAIESDSVQP 66
          |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MAASLLAVLLLLLERGMFSSPSPPPPALLEKVFQYIDLHQDEFVQTLKEWVAIESDSVQP 60

Qy   67 VPRRFQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAEIGSDPTKGT 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VPRRFQELFXMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAEIGSDPTKGT 120

Qy   127 VCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSAFRALEQD 186
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   121 VCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSAFRALEQD 180

Qy   187 LPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAITYGTRGN 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAITYGTRGN 240

Qy   247 SYFMVEVKCRDQDFHSGTFGGIIHEPMADLVALLGSLVDSSGHILVPGIYDEVVPLTEEE 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 SYFMVEVKCRDQDFHSGTFGGIIHEPMADLVALLGSLVDSSGHILVPGIYDEVVPLTEEE 300

Qy   307 INTYKAIHLDLEEYRNSSRVEKFLPDTKEEILMHLWRYPSLSIHGIEGAPDEPGTKTVIP 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 INTYKAIHLDLEEYRNSSRVEKFLPDTKEEILMHLWRYPSLSIHGIEGAPDEPGTKTVIP 360

Qy   367 GRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWIANIDDTQ 426
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWIANIDDTQ 420

Qy   427 YLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHSQNEKINR 486
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 YLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHSQNEKINR 480

Qy   487 WNYIEGTKLFAAFFLEMAQLH 507
          |||||||||||||||||||:|
Db   481 WNYIEGTKLFAAFFLEMAQIH 501

Search completed: February 17, 2005, 17:34:39
Job time : 1 secs

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: February 17, 2005, 17:08:51 ; Search time 0.001 Seconds
                       (without alignments)
                       75.036 Million cell updates/sec

Title:
Perfect score: 2623
Sequence: 1 MDPKLGRMAASLLAVLLLLL................NYIEGTKLFAAFFLEMAQLH 507

Scoring table: BLOSUM62
                    Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 148 residues

Total number of hits satisfying chosen parameters: 1

Database :   us-09-621-976-3957:*

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

                         %
Result   Query
No.   Score Match Length DB ID   Description
------ ----- ----- ------ -- ---
1   607.5 23.2   148   1 us-09-621-976-3957

                              ALIGNMENTS

RESULT 1
us-09-621-976-3957

Query Match   23.2%; Score 607.5; DB 1; Length 148;
Best Local Similarity   96.9%; Pred. No. 0; Mismatches 3;
Matches 124; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy   1 MDPKLGRMAASLLAV-LLLLLERGMFSSPSPSPPALLEKVFQYIDLHQDEFVQTLKEWVAI 59
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1 MDPKLGRMAASLLAVLLLLLLERGMFSSPSPSPPALLEKVFQYIDLHQDEFVQTLKEWVAI 60

Qy   60 ESDSVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQOLPDGQSLPIPPVILAELG 119
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ESDSVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQOLPDGQSLPIPPVILAELX 120

Qy   120 SDPTKGTV 127
          |||| ||
Db   121 SDPTXXTV 128

Search completed: February 17, 2005, 17:08:51
Job time : 0.001 secs

```
OM protein - protein search, using sw model

                                GenCore version 5.1.6
                                Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:              February 17, 2005, 16:21:08 ; Search time 0.001 Seconds
                                                    (without alignments)
                                                    192.153 Million cell updates/sec

Title:          u8-10-036-342-57
Perfect score:  2623
Sequence:       1 MDPKLGRMAASILAVLLLL..........NYIEGTIKLFAAFFLEMAQLH 507

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1 seqs, 379 residues

Total number of hits satisfying chosen parameters:    1

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       ab1108.pir2:*

                Pred. No. is the number of results predicted by chance to have a
                score greater than or equal to the score of the result being printed,
                and is derived by analysis of the total score distribution.


                                        SUMMARIES

          Query                Match
Result    %
No.   Score Match Length DB ID         Description
----- ----- ----- ------ -----         -----------
    1 211.5   8.1  379 1  AB1108      - succinyldiaminop
```

```
C;Genetics:
A;Gene: lmo0265
C;Superfamily: succinyl-diaminopimelate desuccinylase
```

```
Query Match        8.1%;  Score 211.5; DB 1;  Length 379;
Best Local Similarity    23.0%;  Pred. No. 0;
Matches 105; Conservative  75; Mismatches 176; Indels 101; Gaps 19;
```

```
Qy     44 LHQDEFVQTLKEWVAIESDS--VQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQQ 101
             ||| |   || |  | ||     |   |:  || |  ||: | :||   |  |   |:
Db      1 MDDQKKIQILKDLVNIDSTNGHEEQVANYLQKLLAEHGIESEKVQ-------YDL---- 48

Qy    102 LPDCGQSLPIPPVILAELGSDPFKGTVCPYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGR 161
           |:||:| |||| ||  ||||| |  :  | | | ||  |  | ||  || ||| |||||||
Db     49 --DRASL------VSEIGSSNEK-VLAFSGHMDVVDAGDVSKWKFPPFEATEHBGKLYGR 99

Qy    162 GATDNKGPVLAWINAVSAFRALEQDLPVNIKFIIEGMEEAGSVALEELVEKEKDRFSGV 221
           ||||:||| | ||||| ||||  ||| ||| | | |||| | ||||||| |||:|  :|
Db    100 GATDMKSGLAAMVIAMIELHEEKQKLINGKIRLLATVGEEIGELGAEQLTQK--GYADDL 156

Qy    222 DYIVISDNLWISQRKPA---ITYGTRGNSYFMVEVKCRDQDFHSG--TFGILHEPWADL 276
           |  ::|| |    ||     :  || | :  |   | |   | |     |:| |   |
Db    157 DGLIIGE------PSGHRIVYAHKGSINY--TVKSTGKNAHSSMPEFG-------- 196

Qy    277 VALLGSLVDSSGHILVPGIYDEVVPLTEEEINTYKAIHDLEEYRNSSRVEKFL--FDTK 334
           |:  |   ||:  |   |              ||| :  || :    :|::||    |
Db    197 ------VNAIDNLLL--FYNE--------------------------VEKFVKSIDAT 220

Qy    335 EEILMHLWRYPSLSIHGIEGAFDEPGTKTVIPGRVIGKSRSIRLVPHMNVSAVEKQVTRHL 394
             |:|  || || | | | ||        :| | |||| : ||  ||:|   |||  | |
Db    221 NEILGDF------IHNVT-VIDGGNQVNSIPEKAQLQGNIRSIPEMDNETV-KQVLVKI 271

Qy    395 EDVFSKRNSSNKMVVSMTLGLHPWIANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIA 454
            ||  ::|:|:   :  || |  | |:| |  | ||:|     | | |: | | |  ||
Db    272 INKLNKQENVNLELI-FDYDKQPVFSDKNSDLVHIAKSVASDIVKEEIPLLGISGTTDAA 330

Qy    455 KMPQEIVHKSVVLIPLGAVDDGEHSQNEKINRWNYIE 491
            : |:|| |::  :      ||  :|:|:| | :|||
Db    331 EFTK--AKKEFPVIIFGPGNETFHQVNENVSIGNYLE 365
```

```
Search completed: February 17, 2005, 16:21:08
Job time : 0.001 secs
```

RESULT 1
AB1108
 - succinyldiaminopimelate desuccinylase homolog lmo0265 [imported] -
Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AB1108
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.;
Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.;
E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.;
Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.;
Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.;
Jackson, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.;
Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos,
B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes,
N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1108
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-379 <GLA>
A;Cross-references: UNIPROT:Q9ZEY0; GB:NC_003210; PIDN:CAD00792.1;
PID:gi6409630; GSPDB:GN00177
A;Experimental source: strain EGD-e

OM protein - protein search, using sw model

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

Run on: February 17, 2005, 16:26:43 ; Search time 0.001 Seconds

Title: us-10-036-342-57

Perfect score: 2623

Sequence: 1 MDPKLGRMAASLIAVLLLLL..........NYIEGTIKLFAAFFLEMAQLH 507

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 481 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : s56299.pir2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result                   %     Query
No.   Score   Match   Length   DB    ID         Description
---------------------------------------------------------------------
1     1044    39.8    481      1     S56299     - hypothetical pro
```

ALIGNMENTS

RESULT 1

S56299
- hypothetical protein YFR044c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein F006
C;Species: Saccharomyces cerevisiae
C;Date: 02-Sep-1995 #sequence revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: S56299; S62255; S63791
R;Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma,
S. I.; Sasanuma, M.; Tsuchiya, Y.; Soeda, E.; Yokoyama, K.; Yamazaki, M.;
Tashiro, H.; Eki, T.
submitted to the EMBL Data Library, May 1995
A;Description: Analysis of the nucleotide sequence of chromosome VI from
Saccharomyces cerevisiae.
A;Reference number: S56186
A;Accession: S56299
A;Molecule type: DNA
A;Residues: 1-481 <MUR>
A;Cross-references: UNIPROT:P43616; EMBL:D50617; NID:g836685; PIDN:BAA09283.1;
PID:d1009924; PID:g836799; MIPS:YFR044C
C;Accession: S56299; S62255; S63791
R;Murakami, Y.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma,
M.;
submitted to the EMBL Data Library, December 1994
A;Reference number: S62230
A;Accession: S62255
A;Molecule type: DNA
A;Residues: 1-481 <MUW>
A;Cross-references: EMBL:D44597; NID:g871938; PIDN:BAA08010.1; PID:d1008600;
PID:g871943

R;Eki, T.; Naitou, M.; Hagiwara, H.; Ozawa, M.; Sasanuma, S.I.; Sasanuma, M.;
Tsuchiya, Y.; Shibata, T.; Hanaoka, F.; Murakami, Y.;
Yeast 12, 149-167, 1996
A;Title: Analysis of a 36.2 kb DNA sequence including the right telomere of
chromosome VI from Saccharomyces cerevisiae.
A;Reference number: S63791
A;Accession: S63791
A;Molecule type: DNA
A;Residues: 1-481 <EKI>
A;Cross-references: EMBL:D44597; NID:g871938; PIDN:BAA08010.1; PID:d1008600;
PID:g871943
A;Status: nucleic acid sequence not shown; translation not shown
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June
1995
C;Genetics:
A;Cross-references: SGD:S0001940
A;Map position: 6R
A;Note: YFR044c

Query Match 39.8%; Score 1044; DB 1; Length 481;
Best Local Similarity 45.2%; Pred. No. 0;
Matches 212; Conservative 83; Mismatches 168; Indels 6; Gaps 4;

```
Qy    35  LEKVFQYIDLHQDEFVQTLKEWVAIESDSVQPVPRFRQELFRMMAVAADTLQRLGAR-VA  93
              : :|  |||:||||  :||| ||| :  ::::   :::||:|: ||:: || ||| ||
Db     5  LTSVFQKIDSLKPQFFSRLTK--AIQIPAVSSDESLRSKVFDKAKFISEQLSQSGFHDIK  62

Qy    94  SVDMGPQQLP-DGQSLPIPPVILAELGSDPTKGTVCFYGHLDVQPADRGDGWLTDPY--V  150
              |:| ||||| || :||||| :|| ||||| ||| ||| ||:|||| || |  ||  |
Db    63  MVDLGIQPPISTPNLSLPPVTLSRFGSDPSKKTVLVYGHYDVQPAQLEDGWDTEPFKLV  122

Qy   151  LTEVDGKLYGRGATDNKGPVLAWINAVSAFRALEQDLPVNIKFIIEGMEEAGSVALEELV  210
              |: ||| ||||| ||:||| || | | | || ||: |||| |:||||| || ||:||:
Db   123  IDEAKGIMKGRGVTDDTGPLLSWINVVDAFKASGQEFPVNLVTCFEGMEESGSIKLDELI  182

Qy   211  EKEKRDFFSGVDYIVISDNLWISQRKPAITYGTRGNSYPMVEVKCRDQDFHSGTFGGILH  270
              || |  | ||||:|||||| ||: | :|||| |||| :|: | | |||  |:|| |:|
Db   183  KKEANGYFKGVDAVCISDNYWLGTKKVLTYGLRGCNYYQTIIEGPSADLHSGIFGGVVA  242

Qy   271  EPMADLVALLGSLVDSSGHILVPGIYDEVVPLTEEEINTYKAIHDLEEYRNSSRVEKFL  330
              |||::||| || ||  ||:|||| |:|||||:||||:| ||||::||| |||| ||||
Db   243  EPMIDLMQVLGSLVDSKGKILIDGIDEMVAPLTEKEKALYKDIEFSVEELNAATGSKTSL  302

Qy   331  FDTKEEILMHLWRYPSLSISHGIEGAFDEPGTKTVIPGRVIGKFISIRLVPHNVSAVEKQV  390
              |||:||:| ||||||||: || ||: |:||||| :|| || |||| || |||| :|||:|
Db   303  YDKKEDILMHRWRYPSLSIHGVEGAFSAQGAKTVIPAKVFGKFSIRTVPDMDSEKLTSLV  362

Qy   391  TRHLEDVFSKRNSSNKMVVSMTLGLHPWIANIDDTQYLAAKRAIRTVFGTEPDMIRDGST  450
              |  :  :  ::  :|: : :    || | | ||||||| ||  | |:| | ||:| ||
Db   363  QKHCDAKFKSLNSPNKCRTELIHDGAYWVSDPNAQFTPAAKKATKLVYGVPDFTREGGS  422

Qy   451  IPIAKMFQEIVHKSVVLIPLGAVDDGEHSQNEKINRWNYIEGTKLFAAF  499
              |||||:|||||||||||||||||| ||:|| | ||| | ||||| |||
Db   423  IPITLTFQDALNTSVLLLPMGRGDDGAHSINEKLDISNFVGGMKTMAAY  471
```

Search completed: February 17, 2005, 16:26:43
Job time : 0.001 secs

OM protein - protein search, using sw model

Run on: February 17, 2005, 16:29:09 ; Search time 0.001 Seconds
                                     (without alignments)
                                     251.979 Million cell updates/sec

Title: UB-10-036-342-57
Perfect score: 2623
Sequence: 1 MDPKLGRMAASLLAVLLLLL.............NYIEGTKLIFAAFFLEMAQLH 507

Scoring table: BLOSUM62
             Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 497 residues

Total number of hits satisfying chosen parameters: 1

Post-processing: Minimum Match 0%
             Maximum Match 100%
             Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   s43745.pir2:*

             Pred. No. is the number of results predicted by chance to have a
             score greater than or equal to the score of the result being printed,
             and is derived by analysis of the total score distribution.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

SUMMARIES

Result No.	Score	% Query Match Length	DB ID	Description
1	47	1.8	497	1	S43745	- phosphatidylinos

ALIGNMENTS

RESULT 1
S43745
- phosphatidylinositol-phosphatidylcholine transfer protein SEC14 -
yeast (Yarrowia lipolytica).
C;Species: Yarrowia lipolytica, Candida lipolytica
C;Date: 03-May-1994 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
C;Accession: S43745
R;Lopez, M.; Nicaud, J.; Vergnolle, C.; Kader, J.; Bankaitis, V.; Gaillardin, C.
submitted to the EMBL Data Library, July 1993
A;Description: A phospholipid transfer protein is required for dimorphic
transition in the yeast Yarrowia lipolytica.
A;Reference number: S43745
A;Accession: S43745
A;Molecule type: DNA
A;Residues: 1-497 <LOP>
A;Cross-references: UNIPROT:P45816; EMBL:L20972; NID:g311166; PID:g311167
C;Genetics:
A;Gene: SEC14
A;Introns: 6/3; 9/2
F;58-265/Domain: cellular retinaldehyde-binding protein homology <CRB>

Query Match   1.8%; Score 47; DB 1; Length 497;
Best Local Similarity   28.2%; Pred. No. 0;
Matches   40; Conservative   14; Mismatches   54; Indels   34; Gaps   7;

Qy   12 LLAVLLLLLERGMFSSPSPPPALL-EKVFQYIDLHQDEFVQTLKEWVAIESDSVQPVPRF 70
             |||||||||| | | || : | | : | || ||: : | : |

Db   349 LLAVLLLLLELLLPSRLLLPRLLLPRRQGSRSRCCQDRHCPCPHRQDCCTSDQL------F 403

Qy   71 RQELFRMM------------AVAADTLQRLGA--------RVASVDMGPQQLPDG--QSLP- 109
             | | | | | | | | | ||| |

Db   404 PPEMLAWMEQNEKKAEKAAAEAKTKGGAAPQANAFTDEQLAQLNSLPQQVATGGVSQQPLPD 463

Qy   110 ------IPPVI-LAELGSDPTK 124
                        | | | | : |

Db   464 SHVTQGVPTVAKIAELPPVPMR 485

Search completed: February 17, 2005, 16:29:09
Job time : 0.001 secs

```
                                          GenCore version 5.1.6
                                          Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:                    February 17, 2005, 16:33:34 ; Search time 0.001 Seconds
                                                          (without alignments)
                                                          257.049 Million cell updates/sec

Title:
Perfect score:       2623
Sequence:            1 MDPKLGRMAASLLAVLLLLL...........NYIEGTKLFAAFFLEMAQLH 507
Scoring table:       us-10-036-342-57
Scoring table:       BLOSUM62
                     Gapop 10.0 , Gapext 0.5
Searched:            1 seqs, 507 residues

Total number of hits satisfying chosen parameters:    1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :  q6uwk2.uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                    SUMMARIES

Result    Score   Query       Match     DB   ID              Description
No.               Match    Length   DB
-----    -----   -----    ------   ---   -------          -----------
  1      2623    100.0     507      1    Q6UWK2           DPKL1915. Homo sa


                                    ALIGNMENTS

RESULT 1
Q6UWK2
ID   Q6UWK2                    PRELIMINARY;       PRT;    507 AA.
AC   Q6UWK2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   DPKL1915.
GN   ORFNames=UNQ1915;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA   Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA   Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA   Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA   Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA   Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA   Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA   Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.H., Yansura D.,
RA   Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA   Godowski P.;
RT   "The secreted protein discovery initiative (SPDI), a large-scale
RT   effort to identify novel human secreted and transmembrane proteins: a
```

```
RT   bioinformatics assessment.";
RL   Genome Res. 13:2265-2270(2003).
DR   EMBL; AY358756; AAQ89116.1; -.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002933; P:proteolysis and peptidolysis; IEA.
DR   Pfam; PF01546; Peptidase_M20; 1.
```

```
SQ   SEQUENCE   507 AA;  56692 MW;  F3B51A9123C927C0 CRC64;

Query Match             100.0%;   Score 2623;   DB 1;   Length 507;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches  507;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 MDPKLGRMAASLLAVLLLLLERGMFSSPSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDPKLGRMAASLLAVLLLLLERGMFSSPSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAIE 60

Qy         61 SDSVQPVPRFRQELFRMMAVAAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELGS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SDSVQPVPRFRQELFRMMAVAAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELGS 120

Qy        121 DPFKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSAF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DPFKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSAF 180

Qy        181 RALEQDLPVNIKFIIEGMEERAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RALEQDLPVNIKFIIEGMEERAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAIT 240

Qy        241 YGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEVV 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEVV 300

Qy        301 PLTEEEINTYKAIHDLEEYRNSSRVEKFLFDTKEEIIMHLWRYPSLSIHGIEGAFDEPG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PLTEEEINTYKAIHDLEEYRNSSRVEKFLFDTKEEIIMHLWRYPSLSIHGIEGAFDEPG 360

Qy        361 TKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWIA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWIA 420

Qy        421 NIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHSQ 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 NIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHSQ 480

Qy        481 NEKINRWNYIEGTKLFAAFFLEMAQLH 507
              |||||||||||||||||||||||||||
Db        481 NEKINRWNYIEGTKLFAAFFLEMAQLH 507


Search completed: February 17, 2005, 16:33:34
Job time : 0.001 secs
```

OM protein - protein search, using sw model

```
                        GenCore version 5.1.6
                        Copyright (c) 1993 - 2005 Compugen Ltd.


Run on:                 February 17, 2005, 16:35:37

Title:                  February 17, 2005, 16:35:37 ; Search time 0.001 Seconds
                                                       (without alignments)
                                                       257.049 Million cell updates/sec
Perfect score:          2623
Sequence:               US-10-036-342-57
                        1 MDPKLGRMAASLLAVLLLLL.........NYIEGTKLFAAFFLEMAQLH 507

Scoring table:          BLOSUM62
                        Gapop 10.0 , Gapext 0.5

Database :              aaq89116.genpept:*

Post-processing:        Minimum Match 0%
                        Maximum Match 100%
                        Listing first 45 summaries

Minimum DB seq length:  0
Maximum DB seq length:  2000000000

Total number of hits satisfying chosen parameters:       1

Searched:               1 seqs, 507 residues

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

```
Result                  Query
No.     Score  Match  Length  DB  ID         Description
-----   -----  -----  ------  --  --         -----------
1       2623   100.0  507     1   AAQ89116
```

ALIGNMENTS

RESULT 1
AAQ89116

```
Query Match         100.0%;  Score 2623;  DB 1;  Length 507;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches  507;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy     1 MDPKLGRMAASLLAVLLLLLERGMFSSPSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAIE 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MDPKLGRMAASLLAVLLLLLERGMFSSPSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAIE 60

Qy    61 SDSVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELGS 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 SDSVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELGS 120

Qy   121 DPTKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSAF 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 DPTKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSAF 180

Qy   181 RALEQDLPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAIT 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RALEQDLPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAIT 240

Qy   241 YGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEVV 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 YGTRGNSYFMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEVV 300

Qy   301 PLTEEEINTYKAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAPDEPG 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 PLTEEEINTYKAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAPDEPG 360

Qy   361 TKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWIA 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWIA 420

Qy   421 NIDDTQYLAAKRAIRTVFGTEPDMIRDGSTTPIAKMFQEIVHKSVVLIPLGAVDDGEHSQ 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 NIDDTQYLAAKRAIRTVFGTEPDMIRDGSTTPIAKMFQEIVHKSVVLIPLGAVDDGEHSQ 480

Qy   481 NEKINRWNYIEGTKLFAAFFLEMAQLH 507
           ||||||||||||||||||||||||||||
Db   481 NEKINRWNYIEGTKLFAAFFLEMAQLH 507
```

Search completed: February 17, 2005, 16:35:37
Job time : 0.001 secs

```
                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          February 17, 2005, 16:38:39 ; Search time 0.001 Seconds
                                               (without alignments)
                                               257.556 Million cell updates/sec

Title:           us-10-036-342-57
Perfect score:   2623
Sequence:        1 MDPKLGRWAASLLAVLLLLLL..........NYIEGTKLPAFFLEMAQLH 507

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1 seqs, 508 residues

Total number of hits satisfying chosen parameters: 1

Database :       cgl2_human.uniprot_sprot.*
                 Listing first 45 summaries

Post-processing: Minimum Match 0%
                 Maximum Match 100%

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.


                                SUMMARIES

Result                  %
No.    Score  Match  Length  DB   ID           Description
------------------------------------------------------------
  1    2585.5  98.6   508    1   CGL2_HUMAN    Glutamate carboxyp


                                ALIGNMENTS

RESULT 1
CGL2_HUMAN   Q96KN2; Q9BT98;

ID   CGL2_HUMAN     STANDARD;      PRT;   508 AA.
AC   Q96KN2; Q9BT98;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Glutamate carboxypeptidase-like protein 2 precursor (CNDP dipeptidase
DE   1).
GN   Name=CNDP1; Synonyms=CPGL2;
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE OF 199-508 FROM N.A.
RC   TISSUE=Skin;
RA   Chen J.M., Barrett A.J.;
RT   "Cloning and sequencing of a second human homologue of glutamate
RT   carboxypeptidase in peptidase family M20.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the peptidase M20A family.
CC   -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AJ417564; CAD10388.1; -.
DR   EMBL; BC004271; -; NOT_ANNOTATED_CDS.
DR   MEROPS; M20.006; -.
DR   Genew; HGNC:20675; CNDP1.
DR   InterPro; IPR001261; ArgE_dapE.
DR   InterPro; IPR002933; Peptidase_M20.
DR   Pfam; PF01546; Peptidase_M20; 1.
DR   PROSITE; PS00758; ARGE_DAPE_CPG2_1; FALSE_NEG.
DR   PROSITE; PS00759; ARGE_DAPE_CPG2_2; FALSE_NEG.
KW   Carboxypeptidase; Hydrolase; Metalloprotease; Signal.
FT   SIGNAL        1     27       Potential.
FT   CHAIN        28    508       Glutamate carboxypeptidase-like protein
FT   CONFLICT    238    238       L -> P (in Ref. 2).
FT   CONFLICT    273    273       L -> P (in Ref. 2).
SQ   SEQUENCE   508 AA;  56779 MW;  0FDEA991FD8495D CRC64;

Query Match    98.6%;  Score 2585.5;  DB 1;  Length 508;
Best Local Similarity   99.2%;
Matches   504;  Conservative   0;  Mismatches   3;  Indels   1;  Gaps   1;

Qy    1 MDPKLGRWAASLLAV-LILLLERGMFSSPSPSPPPALLEKYFQYIDLHQDEFVQTLKEWVAI 59
         ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    1 MDPKLGRWAASLLAVLLLLLLERGMFSSPSPSPPPALLEKYFQYIDLHQDEFVQTLKEWVAI 60

Qy   60 ESDSVQPVPRPRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELG 119
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ESDSVQPVPRPRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELG 120

Qy  120 SDPTKGTVCYGHLDVQPADRGDGWLTDPYVLTEVDGKLYGRGAIDNKGPVLAWINAVSA 179
         |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  121 SDPTKGTVCFYGHLDVQPADRGDGWLTDPYVLTEVGGKLYGRGAIDNKGPVLAWINAVSA 180

Qy  180 FRALEQDLPVNIKFIIEGMEEAGSVALEELVEKKDRFSGVDYIVISDNIWISQRKPAI 239
         |||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |
Db  181 FRALEQDLPVNIKFIIEGMEEAGSVALEELVEKKDRFSGVDYIVISDNLWISQRKLAI 240

Qy  240 TYGTRGNSYPWVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEV 299
         ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  241 TYGTRGNSYPWVEVKCRDQDFHSGTFGGILHELMADLVALLGSLVDSSGHILVPGIYDEV 300

Qy  300 VPLTEEEINTYKAIHDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAFDEP 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 VPLTEEEINTYKAIHDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAFDEP 360
```

```
Qy  360 GTKTVIPGRVIGKFSIRLVPHNNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWI 419
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GTKTVIPGRVIGKFSIRLVPHNNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWI 420

Qy  420 ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHS 479
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 ANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHS 480

Qy  480 QNEKINRWNYIEGTKLFAAFFLEMAQLH 507
        ||||||||||||||||||||||||||||
Db  481 QNEKINRWNYIEGTKLFAAFFLEMAQLH 508
```

Search completed: February 17, 2005, 16:38:39
Job time : 0.001 secs

```
OM protein - protein search, using sw model

                                        GenCore version 5.1.6
                                        Copyright (c) 1993 - 2005 Compugen Ltd.

· Run on:                February 17, 2005, 16:40:08 ; Search time 0.001 Seconds
                                                        (without alignments)
                                                        86.697 Million cell updates/sec

  Title:                 us-10-036-342-57
  Perfect score:         2623
  Sequence:              1 MDPKLGRWAASLLAVLLLLL.........NYIBGTKLFAAFFLEMAQLH 507

  Scoring table:         BLOSUM62
                         Gapop 10.0 , Gapext 0.5

  Searched:              1 seqs, 171 residues

  Total number of hits satisfying chosen parameters:    1

  Minimum DB seq length: 0
  Maximum DB seq length: 2000000000

  Post-processing:       Minimum Match 0%
                         Maximum Match 100%
                         Listing first 45 summaries

  Database :  q6znd4.uniprot_trembl:*

            Pred. No. is the number of results predicted by chance to have a
            score greater than or equal to the score of the result being printed,
            and is derived by analysis of the total score distribution.

                                          SUMMARIES

  Result                 %      Query
  No.     Score    Match    Length DB  ID              Description
  ------  -----  -------  ------ --  --------------  --------------------
    1      766    29.2    171   1   Q6ZND4          Hypothetical prote


                                         ALIGNMENTS

RESULT 1
ID    Q6ZND4
AC    Q6ZND4;
DT    05-JUL-2004 (TrEMBLrel. 27, Created)
DT    05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT    05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE    Hypothetical protein FLJ16195.
OS    Homo sapiens (Human).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX    NCBI_TaxID=9606;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Thalamus;
RA    Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
RA    Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
RA    Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA    Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA    Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA    Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA    Masuho Y., Nagai K., Isogai T.;
RL    Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR    EMBL; AK131263; BAD18441.1; -.
SQ    SEQUENCE  171 AA;  18880 MW;  20816D8F214688 CRC64;

Query Match            29.2%;  Score 766;  DB 1;  Length 171;
```

```
Best local Similarity  100.0%;  Pred. No. 0;
Matches 148;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      339  MHLWRYPSLSIHGIEGAFDEBPGTKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVF 398
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1  MHLWRYPSLSIHGIEGAFDEBPGTKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVF 60

Qy      399  SKRNSSNKMVVSMTLGLHPWIANIDDTQYLAAKRAIRTVFGTETEPDMIRDGSTIPIAXMFQ 458
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61  SKRNSSNKMVVSMTLGLHPWIANIDDTQYLAAKRAIRTVFGTETEPDMIRDGSTIPIAXMFQ 120

Qy      459  EIVHKSVVLIPLGAVDDGEHSQNEKINR 486
             ||||||||||||||||||||||||||||
Db      121  EIVHKSVVLIPLGAVDDGEHSQNEKINR 148

Search completed: February 17, 2005, 16:40:09
Job time : 0.001 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 17:06:12 ; Search time 0.001 Seconds
                                                     (without alignments)
                                                     86.697 Million cell updates/sec

Title: us-10-036-342-57
Perfect score: 2623
Sequence: 1 MDPKLGRMAASLLAVLLLL.........NYIEGTKLFAAFPLEMAQLH 507

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 171 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   bad18441.genpept:*

              Pred. No. is the number of results predicted by chance to have a
              score greater than or equal to the score of the result being printed,
              and is derived by analysis of the total score distribution.

                                       SUMMARIES

Result   Query
No.   Score   Match Length DB ID   Description
---- -----   ----- ------ -----   -----------
1   766   29.2   171   1   BAD18441

                                      ALIGNMENTS

RESULT 1
BAD18441

Query Match   29.2%;   Score 766; DB 1; Length 171;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   339 MHLWRYPSLSLIHGIEGAFDEPGTKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVF 398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MHLWRYPSLSLIHGIEGAFDEPGTKTVIPGRVIGKFSIRLVPHMNVSAVEKQVTRHLEDVF 60

Qy   399 SKRNSSNKMVVSMTLGLHPWIANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQ 458
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 SKRNSSNKMVVSMTLGLHPWIANIDDTQYLAAKRAIRTVFGTEPDMIRDGSTIPIAKMFQ 120

Qy   459 EIVHKSVVLIPLGAVDDGEHSQNEKINR 486
          ||||||||||||||||||||||||||||
Db   121 EIVHKSVVLIPLGAVDDGEHSQNEKINR 148

Search completed: February 17, 2005, 17:06:12
Job time : 0.001 secs